Diagonal representation for the transfer-matrix method for obtaining electronic energy levels in layered semiconductor heterostructures

B. Chen, M. Lazzouni, and L. R. Ram-Mohan
Department of Physics, Worcester Polytechnic Institute, Worcester, Massachusetts 01609
(Received 18 March 1991; revised manuscript received 8 August 1991)

We develop a multiband $\mathbf{k}\cdot\mathbf{P}$ transfer-matrix algorithm by defining a "diagonal representation" that provides a unified way of calculating energy levels and wave functions for superlattices as well as for quantum wells with arbitrarily shaped band-edge profiles. Numerical results of transfer-matrix calculations are presented for specific heterostructures. We show that, contrary to expectations, the various versions of the boundary conditions used in the literature lead to results for the subband energies that agree with one another reasonably well for the wide-band-gap materials.

I. INTRODUCTION

The envelope-function approximation (EFA) was introduced several years ago for the analysis of wave functions and energy levels in quantum heterostructures.[1,2] The basic ideas entering this model parallel those of effective-mass theory,[3] in the framework of the $\mathbf{k}\cdot\mathbf{P}$ models used in the theory of electronic band structure in bulk semiconductors.[4] With its foundation in the $\mathbf{k}\cdot\mathbf{P}$ model, the EFA is a phenomenological approach that takes advantage of the experimentally available[5] band parameters of the constituent semiconductors of the heterostructures as input. In the case of the conduction-band energy levels in quantum-well (QW) and superlattice (SL) heterostructures of III-V and II-VI constituent materials with relatively large band gaps, a further simplification obtains in that a one-band analysis (analogous to the Kronig-Penney model[6]) suffices. It is this simplicity and the empirical nature of this theory which has led to its widespread use, especially for $GaAs/Al_xGa_{1-x}As$ heterostructures. These advantages have made the model amenable to the inclusion of band nonparabolicity effects,[7] to its extension to multiband treatments, to the inclusion of effects of built-in strain, and to the presence of external magnetic and electric fields.[1,8–15] In a multiband formalism, the EFA has been extended also to heterostructures with narrow-band-gap materials, such as HgTe/CdTe superlattices, in the presence of an external magnetic field.[16–18]

The purpose of this paper is to present a unified way of obtaining the solution of the multiband differential equations for QW's as well as for SL's. Earlier, one of us[16] had provided a compact and efficient algorithm for the numerical evaluation of the electronic energy bands only in the superlattices using the transfer-matrix method (TMM). Here we show that the use of a "diagonal representation" of the transfer matrix for each heterolayer leads to an analogous algorithm for the energy levels in QW's of any band-edge profile. It is well known[14,11] that the $\mathbf{k}\cdot\mathbf{P}$ Hamiltonian for bulk semiconductors has spurious solutions with values for the wave vector k which lie outside of the first Brillouin zone; this is true for any given value of the energy E for models with eight or more bands. (This is discussed further in Sec. II.) The diagonal representation permits the identification of these spurious states associated with the eight-band $\mathbf{k}\cdot\mathbf{P}$ model and allows their elimination for each layer, resulting in a numerically stable, compact algorithm for the eigenstates and eigenenergies. It should be recalled that the exponentiation of matrices, in general, is fraught with numerical pitfalls;[19] these are no longer an issue when the diagonal representation is employed and the contributions of the spurious states are eliminated in each layer.

The variety of developments in the theory of electronic energy levels in layered semiconductors has inevitably led to various forms of the boundary conditions being used in the solution of Schrödinger's equation for the energy levels and wave functions. The different implementations of the theory also vary, so that there is a clear need to take stock of the variations in these approaches. We compare the numerical results obtained for QW energy levels when different boundary conditions are used.

The paper is arranged as follows. In Sec. II, the EFA is used to define the $\mathbf{k}\cdot\mathbf{P}$ Hamiltonian and the Schrödinger's equations satisfied by the envelope functions. These second-order differential equations are solved by writing them as coupled first-order differential equations. We define the diagonal representation of the transfer-matrix solutions of these equations and demonstrate the advantages of this representation.

In Sec. III, the theory is developed for application to superlattices and to QW's of arbitrary band-edge profile. These modifications of the TMM lead to a unified way of treating the QW and SL problems. The diagonal representation allows us to implement two specific boundary conditions used earlier. The numerical results are presented for $GaAs/Al_xGa_{1-x}As$ and $In_{1x}Ga_xAs_yP_{1-y}/InP$ lattice-matched QW's of different well shapes under the different boundary conditions. The discussion in Sec. IV provides comparison of the present approach with earlier ones.

II. THE DIAGONAL REPRESENTATION FOR TRANSFER MATRICES

A. Preliminary remarks

We consider planar layered semiconductor heterostructures with the planes perpendicular to the growth direction z. The layers are taken to be composed of compound III-V or II-VI semiconductors with their conduction- and valence-band edges located at the Γ point in the Brillouin zone. The periodic parts of the Bloch functions, $u_{j,\mathbf{k}=0}(\mathbf{r})$ with j being the band index, at the band edges are assumed not to differ much as we traverse layer interfaces. We shall suppose that the original bulk crystal translational symmetry is maintained in the transverse direction.

We shall be interested in the zone-center bulk band structure of the constituent semconductors, within the spirit of the $\mathbf{k}\cdot\mathbf{P}$ model. The usual eight-band model consists of the Γ_6 conduction band (c), the Γ_8 heavy-hole (hh) and light-hole (lh) bands, and the Γ_7 spin-orbit split-off (SO) band, with their spin degeneracies. In the following, we limit ourselves to the case with no external electric or magnetic fields or built-in strain in the layers. (The following considerations hold for the more general case with external perturbations except that the Kramer's degeneracy of the bands gets lifted, and the dimensions of the matrices are larger. The latter would be true also in the presence of strain). With this degeneracy and with the in-plane wave vector $(k_x,k_y)=0$, the problem reduces to a three-band model, with the hh band factoring out. The problem of solving for the envelope functions of the constituent layers, within EFA, then reduces to the solution of a set of three simultaneous second-order differential equations for the envelope functions. We have

$$H_{ij}(\mathbf{k}_\perp,k_z)f_j(z)=Ef_i(z)\ , \tag{1}$$

where k_z has to be replaced by the differential operator $-i\partial/\partial z$, and $\mathbf{k}_\perp=(k_x,k_y)$ is the in-plane wave vector which is set to zero in the following. The eigenvalues of the 3×3 matrix are given by the secular equation (in atomic units):

$$\begin{vmatrix} E_c+(F+\frac{1}{2}k_z^2)-E & -\sqrt{E_P/3}k_z & -\sqrt{E_P/6}k_z \\ -\sqrt{E_p/3}k_z & E_v-\frac{1}{2}(\gamma_1+2\gamma_2)k_z^2-E & -\sqrt{2}\gamma_2k_z^2 \\ -\sqrt{E_P/6}k_z & -\sqrt{2}\gamma_2k_z^2 & E_s-\frac{1}{2}\gamma_1k_z^2-E \end{vmatrix}=0\ . \tag{2}$$

Here E_c, E_v, and E_s are the band-edge energies of the conduction, lh, and SO bands. The three coupled second-order differential equations represented by Eq. (2) can be written as

$$(-\mathcal{A}_{ab}\partial^2/\partial z^2-i\mathcal{B}_{ab}\partial/\partial z+\mathcal{C}_{ab})f_b(z)=Ef_a(z)\ . \tag{3}$$

The matrix coefficients $\underline{\mathcal{A}}$, $\underline{\mathcal{B}}$, and $\underline{\mathcal{C}}$ in Eq. (3) are assumed to be constant in each layer. In a heterostructure the differences in the band-edge energies provide the confining potentials experienced by the carriers.

Equation (2) is solved by writing it out as six coupled first-order differential equations. We begin by defining

$$\Phi(z)=\begin{bmatrix} f(z) \\ f'(z) \end{bmatrix}\ , \tag{4}$$

where the envelope functions $f(z)$ have three components. The differential equation satisfied by $\Phi(z)$ is

$$\begin{aligned}\Phi'(z)&=\begin{bmatrix} \underline{0} & \underline{I} \\ \underline{\mathcal{A}}^{-1}(\underline{\mathcal{C}}-\underline{I}E) & -i\underline{\mathcal{A}}^{-1}\underline{\mathcal{B}} \end{bmatrix}\Phi(z) \\ &\equiv\underline{\Lambda}\Phi(z)\ ,\end{aligned} \tag{5}$$

with the unique solution given by

$$\begin{aligned}\Phi(z)&=\exp(\underline{\Lambda}z)\Phi(0) \\ &=\underline{T}(z)\Phi(0)\ .\end{aligned} \tag{6}$$

Here $\underline{T}(z)=\exp(\Lambda z)$ is called the transfer matrix. Our notation differs slightly from the one used in Ref. 16.

For a given energy E, the TMM functions $\Phi(z)$ can be constructed with the envelope functions $f(z)$ of the form

$$f(z)\sim\sum_a C_a e^{ik_{z,a}z}\ , \tag{7}$$

where k_z is in general complex. This is because they are solutions of second-order differential equations with constant coefficients. The allowed values of k_z are determined by substituting $f(z)$ in the differential equation [Eq. (5)] for $\Phi(z)$. Substituting Eq. (7) into Eq. (5) we obtain

$$\Lambda_{\alpha,\beta}\Phi_\beta(z)=ik_{z,\alpha}\Phi_\alpha(z)\quad(\alpha=1,\ldots,6)\ . \tag{8}$$

Equation (8) is an eigenvalue equation, with the eigenvalues of the $\underline{\Lambda}$ matrix being the general wave vectors ik_z. The components of $f(z)$ are then expressible as *linear combinations* of the exponential functions $\exp(ik_{z,\alpha}z)$ using the eigenvalues $ik_{z,\alpha}$ of the matrix $\underline{\Lambda}$. Note that Eq. (8), which yields the allowed values of $k_{z,\alpha}$, is an eigencondition for k_z; this may be compared with other treatments[9] where a search is made to determine the zeros of the determinant (2) to obtain the allowed values of k_z.

Let us now consider the transformation matrix $\underline{\mathcal{P}}$, which diagonalizes the 6×6 complex general matrix $\underline{\Lambda}$ under a similarity transformation. The eigenvalues of $\underline{\Lambda}$ are

$$\underline{\mathcal{P}}\,\underline{\Lambda}\,\underline{\mathcal{P}}^{-1}=i\underline{k}=ik_{\alpha}\delta_{\alpha\beta}\ ,\quad \alpha,\beta=1,\ldots,6\ . \tag{9}$$

The eigenfunctions $\xi(z)$ given by

$$\xi_{\alpha}(z)=\mathcal{P}_{\alpha\beta}\Phi_{\beta}(z) \tag{10}$$

are explicitly the exponential functions $\exp(ik_{z,\alpha}z)$ and as such also satisfy Eq. (5). The above transformation $\mathcal{P}$ has thus transformed the envelope functions $\Phi(z)$ into pure exponentials. *We define the diagonal representation for the transfer matrix to be the one in which the matrix $\underline{\Lambda}$ is diagonal.* This diagonal representation has a number of advantages, as we demonstrate below.

B. Interface boundary conditions

The boundary conditions at interfaces in planar heterostructures are (i) the continuity of the envelope functions $f(z)$ at the interface [we have already imposed the continuity of the periodic part $u_{j,\mathbf{k}=0}(\mathbf{r})$ of the Bloch functions], and (ii) the continuity of the probability current across the interface.

To derive the probability current, we begin by requiring that the envelope functions $f(z)$ be normalized. The time derivative of $[\int dz\, f^{\dagger}(z)f(z)]$ being zero, we use Eq. (3) to derive the form of the conserved current:

$$j(z)=f^{\dagger}(z)\left[-\underline{\mathcal{A}}f'(z)-\frac{i}{2}\underline{\mathcal{B}}f(z)\right]+\left[f^{\dagger\prime}(z)\underline{\mathcal{A}}-f^{\dagger}(z)\frac{i}{2}\underline{\mathcal{B}}\right]f(z)\ , \tag{11}$$

where the Hermiticity of $\underline{\mathcal{A}}$, $\underline{\mathcal{B}}$, and $\underline{\mathcal{C}}$ has been used.

Thus the current continuity condition across interfaces is given by the continuity of

$$\left[-\underline{\mathcal{A}}f'(z)-\frac{i}{2}\underline{\mathcal{B}}f(z)\right]\ . \tag{12}$$

In the following we refer to the boundary condition (12) as BC1.

In the earlier presentation[16] of TMM, the current continuity condition was replaced by the requirement that the Schrödinger equation [Eq. (3)] be valid throughout the layers and also across the interfaces. This condition led to the continuity of

$$[-\underline{\mathcal{A}}f'(z)-i\underline{\mathcal{B}}f(z)]\ . \tag{13}$$

The discrepancy of a factor of $\frac{1}{2}$ in the coefficient of the $\underline{\mathcal{B}}$ term arises because the corresponding term $-\underline{\mathcal{B}}f'(z)$ in Eq. (3) had not been symmetrized. The discrepancy in the two procedures disappears when the symmetrization is implemented. In the following we refer to the boundary condition (13) as BC2. We compared the effect of imposing BC1 and BC2 on the energy levels in QW's with different band-edge profiles. The energies obtained using BC1 and BC2 agree to within five significant figures. We discuss the reasons for this result in Sec. IV.

If we define $\sigma=\frac{1}{2}$ for BC1 and $\sigma=1$ for BC2 we can represent the interface boundary conditions on the TMM functions $\Phi(z)$ by

$$\begin{bmatrix}\underline{I} & \underline{0}\\ -i\sigma\underline{\mathcal{B}} & -\underline{\mathcal{A}}\end{bmatrix}\Phi(z)|_{L}=\begin{bmatrix}\underline{I} & \underline{0}\\ -i\sigma\underline{B} & -\underline{A}\end{bmatrix}\Phi(z)|_{R}\ , \tag{14}$$

which can be written compactly as

$$\Phi(z)|_{R}=\underline{\mathcal{J}}_{RL}\Phi(z)|_{L}\ , \tag{15}$$

in terms of an interface transfer matrix $\underline{\mathcal{J}}$.

C. Spurious states in the k·P model

The allowed values of k are determined for each given energy E by solving the linear eigenvalue problem, Eq. (8). The original Hamiltonian in Eq. (1) is Hermitian, and the secular equation [Eq. (2)] gives a determinant which is a function of k^2. Hence for every eigenvalue k we must have other eigenvalues $-k$, k^*, and $-k^*$ as allowed solutions.[14] It is found that any adequate treatment of the coupling of both electrons, hh, lh, and SO states by $\mathbf{k}\cdot\mathbf{P}$ theory inevitably leads to the appearance of spurious branches in the dispersion relation.[11] This is because the $\mathbf{k}\cdot\mathbf{P}$ approach is basically a perturbative one, with a finite and incomplete basis set, so that it is impossible for $E(\mathbf{k})$ to be a periodic function of $\mathbf{k}$ when $\mathbf{k}$ moves past the first Brillouin zone. Spurious branches of the dispersion relation appear outside the first Brillouin zone. These branches should be periodic over the entire Brillouin zone, but are found not to be so. Such $E(\mathbf{k})$ branches typically have very large values of $|\mathbf{k}|$ which correspond to very rapidly rising or falling exponential states of highly oscillatory wave functions within the unit cell. Thus states belonging to such k values do not carry any physical significance. They are responsible for occasional numerical instability encountered in the original formulation of TMM.[16] In the context of constructing electronic states in quantum heterostructures, these states have to be eliminated also in order to satisfy the proper boundary conditions.

In our numerical implementation of the TMM algorithm we sort the k eigenvalues and construct a projection matrix $\underline{\mathcal{S}}$, which eliminates the components corresponding to the spurious states in the diagonal representation. Let the spurious states occur for the $\alpha=\mu_1$ and μ_2, say. Then we define

$$\mathcal{S}_{\alpha\beta}=\begin{cases}1 & \text{for } \alpha=\beta\ ,\ \alpha\neq\mu_1,\mu_2\\ 0 & \text{otherwise}\ .\end{cases} \tag{16}$$

We then have the allowed states in the diagonal representation given by $\mathcal{S}_{\alpha\beta}\xi_{\beta}$. Such an elimination of the spurious states is performed in each layer in the heterostructure. This ensures a much better computational stability in the TMM algorithm than anticipated earlier.[16]

Given this need to eliminate the spurious states, it is useful to formulate both the SL and the QW problems in terms of the diagonal representation. We first need to redefine the transformation matrix $\underline{\mathcal{P}}$ which diagonalizes the matrix $\underline{\Lambda}$. The matrix $\underline{\mathcal{P}}$ is made up of the eigenfunctions of $\underline{\Lambda}$ arranged column by column. We therefore set to zero the elements associated with columns (and rows) μ_1 and μ_2. In order to calculate the inverse transforma-

tion matrix $\underline{\mathcal{P}}^{-1}$ we reinsert unit diagonal elements for $\alpha=\beta=(\mu_1,\mu_2)$ in $\underline{\mathcal{P}}$, in which again we set the corresponding rows and columns to zero once the inversion is performed. This procedure of inserting a unit matrix element in the appropriate diagonal elements associated with the spurious states preserves the original dimensions of the matrices in the numerical analysis. We label these transformation matrices with primes. The interface matrices $\underline{\mathcal{I}}$ are also expressed in the diagonal representation. For example, for the interface between layers labeled 1 and 2 we have

$$\underline{\mathcal{I}}'_{12}=\underline{\mathcal{P}}'_1\underline{\mathcal{I}}_{12}\underline{\mathcal{P}}'^{-1}_2 . \tag{17}$$

In Sec. III we apply the above considerations to the SL and the QW geometries.

III. THE TRANSFER-MATRIX METHOD FOR SUPERLATTICES AND QUANTUM WELLS

A. Superlattices

We consider a bilayer superlattice made up of layers of thickness d_1 and d_2, with $d=d_1+d_2$. We have to employ Bloch's periodicity condition on the wave functions, given by

$$\Phi(d)=e^{iqd}\Phi(0) , \tag{18}$$

in order to obtain the correct solutions for the superlattice states and energy bands. Here q is the wave vector of the extended states in the z direction. On the other hand, the wave function at $z=d$ can be obtained from the one at $z=0$ by using transfer matrices for each layer and interface matrices to connect up the wave functions across interfaces. We are thus led to the eigencondition

$$\underline{\mathcal{I}}_{12}\underline{T}_2(d_2)\underline{\mathcal{I}}_{21}\underline{T}_1(d_1)\Phi(0^+)=e^{iqd}\Phi(0^+) . \tag{19}$$

We use the diagonal representation to write Eq. (19) in the form

$$(\underline{\mathcal{P}}'_1\underline{\mathcal{I}}_{12}\underline{\mathcal{P}}'^{-1}_2)e^{ik_2d_2}(\underline{\mathcal{P}}'_2\,\underline{\mathcal{I}}_{21}\mathcal{P}'^{-1}_1)e^{ik_1d_1}\xi(0^+)=e^{iqd}\xi(0^+)$$

or

$$\underline{\mathcal{I}}'_{12}e^{ik_2d_2}\underline{\mathcal{I}}'_{21}e^{ik_1d_1}\xi(0^+)=e^{iqd}\xi(0^+) . \tag{20}$$

Equation (20) is an eigenvalue equation for the function $\xi(0)$. The exponentials $e^{ik_id_i}$ are diagonal matrices with the diagonal elements being given by $e^{ik_{i,\alpha}d_i}$. The energy E, which is an input variable in the differential equation, is varied until the above eigenvalue condition is satisfied. This gives a unique value for the dispersion function $E(q)$. The advantage of the present approach is that, with the spurious roots absent, the transfer matrices are numerically more stable than in the treatment presented in Ref. 16.

B. Quantum wells

Let the quantum well consist of a left barrier region of material labeled by 1, a well region of thickness d of material 2, and a right barrier region of material 3. The barrier regions are considered to be of infinite thickness.

The transfer matrices relate the wave function $\Phi(d/2+0^+)$ in the right barrier to the wave function $\Phi(-d/2-0^+)$ in the left barrier by

$$\Phi(d/2+0^+)=(\mathcal{I}_{32}e^{\Lambda_2d}\underline{\mathcal{I}}_{21})\Phi(-d/2-0^+) . \tag{21}$$

In the diagonal representation Eq. (21) is expressed as

$$\xi(d/2+0^+)=\underline{\mathcal{I}}'_{32}e^{ik_2d}\underline{\mathcal{I}}'_{21}\xi(-d/2-0^-) . \tag{22}$$

Here the boundary conditions at $z=\pm\infty$ require that the wave functions fall off exponentially as $|z|\sim\infty$. We now implement these boundary conditions.

The components of ξ at $z=\pm d/2$ consist of coefficients of exponential functions e^{ikz} with k having both real and imaginary values. The solutions with real values of k_α lead to trigonometric solutions which have to be discarded in the barriers on either side of the quantum well. Furthermore, the states with imaginary values of k_α (call them κ_α) in $\xi(d/2)$ [and in $\xi(-d/2)$] which lead to exponentially increasing solutions with increasing (decreasing) z have to be eliminated. The wave-vector values are thus sorted out first to eliminate the spurious states, then to separate out and discard the ones with real values of k_α, and finally to discard the ones with the improper sign for the wave-vector values which are purely imaginary.

Let us define diagonal projection matrices $\underline{\mathcal{U}}_R$ and $\underline{\mathcal{U}}_L$ which act on $\xi(\pm d/2)$ to eliminate components of $\xi(z)$ which are not evanescent solutions in the two barrier regions, respectively. We can then write

$$\begin{aligned}\underline{\mathcal{U}}_R\xi&=\underline{\mathcal{I}}'_{32}e^{ik_2d}\underline{\mathcal{I}}'_{21}\underline{\mathcal{U}}_L\xi\\&\equiv\underline{\mathcal{Q}}\xi .\end{aligned} \tag{23}$$

Equation (23) is a set of linear equations for the nonzero components of the envelope function amplitudes in the barriers just to the left and the right of the well region. The surviving components in the left amplitude vector, after the projection using $\underline{\mathcal{U}}_L$, do not overlap the ones from the right, after the projection by $\underline{\mathcal{U}}_R$, in the diagonal representation since they correspond to $\pm\kappa_\alpha$ wave vectors, respectively. Hence the unknown components of $\xi(\pm d/2)$ can be combined together and written as a single unknown multicomponent amplitude ξ. [This has been done in Eq. (23), above).] Also, the components corresponding to spurious states factor out. We can therefore write Eq. (23) as a set of homogeneous equations for the components of ξ with a nontrivial solution when the determinant of the coefficients $\det(\underline{\mathcal{U}}_R-\underline{\mathcal{Q}})$ becomes zero. We then scan in energy to search for those energies for which the above determinant is zero in order to obtain the allowed values of energy for bound states. The wave-function amplitudes from such energy are then solved for from the homogeneous equations (23), and separated out again into the wave-function amplitudes just to the left or the right of the quantum well. The full coordinate dependence of the wave functions is then given by exponential functions in the diagonal representation.

For completeness, we note that the above analysis for the allowed energy levels can be cast into the form of an

eigenvalue problem. In Eq. (23), $\underline{\mathcal{U}}_R$ has 1 along the diagonal for the components of ξ which correspond to $-\kappa_\alpha$ and zeros for the other diagonals. Let $\underline{\mathcal{U}}_R^c = \underline{I} - \underline{\mathcal{U}}_R$. Adding $\mathcal{U}_R^c \xi$ to both sides leads to the equation

$$(\underline{Q} + \underline{\mathcal{U}}_R^c)\xi = \underline{I}\xi \ . \tag{24}$$

Equation (24) is thus an eigenvalue problem for ξ with eigenvalue 1, so that the method may be viewed as being parallel to the one for SL's. Again a range of energies is searched with the allowed energy being the one for which the eigenvalue is unity.

IV. APPLICATIONS AND CONCLUDING REMARKS

In this section we investigate numerically specific heterostructures in order to demonstrate the efficacy of the diagonal TMM. We also use the ability of adapting alternative boundary conditions inherent in the diagonal TMM in order to explore a somewhat confusing issue on boundary conditions which has also been considered by Burt.[12] Here we demonstrate the effect of the different boundary conditions numerically. We use both boundary conditions, mentioned in Sec. II, in our calculations of several examples of QW's, and in addition, we include a third set of boundary conditions,[12] referred to as BC3, which are the continuity of

$$f(z) \ , \quad f'(z) \ , \tag{25}$$

across interfaces. The introduction of BC3 helps us making the comparisons more obvious.

In the first example, the subband energies are calculated for the c, hh, and lh states in a GaAs/$Ga_{0.63}Al_{0.37}As$ rectangular QW of 100 Å width. The input band parameters are chosen from Ref. 20 so that not only can we compare the results from the different boundary conditions, namely BC1, BC2, and BC3, but also compare our modified TMM with the original TMM in Ref. 16. The results are listed in Table I. The differences associated with the various BC's are remarkably small. We return to this issue below. In the next example, we consider

TABLE I. Energy levels for the conduction electron (c), light-hole (lh), and heavy-hole (hh) subbands (in meV) for a GaAs/$Ga_{0.63}Al_{0.37}As$ rectangular well of 100 Å width for the three boundary condition (BC) considered in the text. The results are compared with the calculations of Ref. 20 in column 5.

Band	BC1	BC2	BC3	Ref. 20
c	32.53	32.53	32.57	32.53
	122.05	122.05	122.07	122.05
	249.92	242.92	242.85	242.91
lh	22.70	22.70	22.24	22.69
	82.58	82.58	81.60	82.56
	158.10	158.10	157.67	158.09
hh	7.52	7.52	7.57	
	29.92	29.92	30.09	
	66.61	66.61	66.92	
	116.03	116.03	116.43	
	172.42	172.42	172.61	

lattice-matched $GA_{0.47}In_{0.53}As$/InP QW's. The exciton transitions in such QW's have been studied both experimentally and theoretically.[21] The valence-band parameters of $Ga_{0.47}In_{0.53}As$ are not accurately determined. The only experiment on bulk $Ga_{0.47}In_{0.53}As$ for the determination of effective masses of heavy holes and light holes was done by Alavi and Aggarwal.[22] There is a somewhat large uncertainty in the effective mass of heavy holes in the reported value. We shall therefore use the best-known experimental valence-band parameters of GaAs and InAs to calculate the modified Luttinger parameters γ_1 and γ_2, and employ linear interpolation to obtain the valence-band parameters for $Ga_{1-x}In_xAs$. The use of linear interpolation on modified Luttinger parameters rather than on the effective masses of the heavy and light holes can be supported by the following arguments: (i) The experimentally measured band gap of $Ga_{0.47}In_{0.53}As$ is about 10% smaller than that obtained from linear interpolation between GaAs and InAs.[21] Also, the experimental spin-orbit split-off energy,[23] which agrees with the value obtained by Pearsall,[24] is about 10% smaller than that obtained from linear interpolation. This suggests that the bowing effect is important in this material. Thus a direct linear interpolation to determine the effective mass of light holes is not expected to yield an accurate estimate. (ii) The modified Luttinger parameters only contain the interaction from remote conduction bands. Thus they should be less sensitive to change in composition. In fact, the effective mass of light holes obtained from our method is in very close agreement with that determined by Alavi and Aggarwal.[22] The input band parameters are listed in Table II. The calculated subband energies are listed in Table III.

The results displayed in Table I from the three boundary conditions reveal that they play a minor role in the determination of the subband energy edges for the case of a 100-Å GaAs/$Ga_{0.63}Al_{0.37}As$ where the differences between the results obtained from BC1 and BC2 are less than 0.01 meV. Though BC3 looks quite different from BC1 or BC2, the results obtained from BC3 are also no more than a few meV away from those obtained from either BC1 or BC2. And so is the case of $Ga_{0.47}In_{0.53}As$/InP with about a 0.01 to 0.03 meV difference between the results obtained from BC1 and BC2 (Table III); on the other hand, the results from BC3 are off by tens of meV for valence energy levels in the narrower wells with better congruence of values of lower energy levels in wider wells.

These trends can be explained as follows. First, the difference in BC1 and BC2 arises from the treatment of the momentum matrix elements. These are off-diagonal terms in the Hamiltonian. As has been emphasized by Bastard,[1] the momentum matrix elements $\underline{P}$ in all the III-V and II-VI semiconductors are comparable. If we assume the momentum matrix elements are the same for the well and barrier materials, then the difference between BC1 and BC2 disappears. Under the same assumption, the only difference between BC2 and BC3 would come from terms containing the modified Luttinger parameters. Since the modified Luttinger parameters are due to the interaction between valence bands and

TABLE II. Material parameters for InP and $Ga_{0.47}In_{0.53}As$ used in the calculation. The conduction-band offset between them is taken[a] as 0.257 eV. The γ parameters are the modified Luttinger parameters.

Parameter	InP	GaAs	InAs	$Ga_{0.47}In_{0.53}As$
E_g (eV)	1.425[b]	1.518[b]	0.418[b]	0.812[b]
Δ (eV)	0.108[b]	0.340[b]	0.370[b]	0.330[c]
m_e (m_0)	0.079[b]	0.0665[b]	0.023[b]	0.041[c]
m_{lh} (m_0)	0.121[b]	0.091[d]	0.027[b]	0.051[e,f]
m_{hh} (m_0)	0.560[b]	0.34[d]	0.4[b]	0.370[e,f]
E_p (eV)	17.02	22.69	21.05	21.00
F	0.0	0.0	0.0	0.0
γ_1	1.045	1.982	2.982	2.514[f]
γ_2	−0.370	−0.480	0.241	−0.106[f]

[a]Reference 27.
[b]Reference 5.
[c]Reference 24.
[d]Reference 29.
[e]Reference 22.
[f]Obtained by linear interpolation between the modified Luttinger parameters of GaAs and InAs (Ref. 5) and in very close agreement with Ref. 22.

remote conduction bands, their effect on the determination of the sub-band-edge energies is much less important. Therefore the differences in all the three boundary conditions is insignificant, particularly for a wide quantum well, in this limit. For a narrower well the absence of the "mass-derivative" factors in the continuity conditions of BC3 leads to substantially higher valence-band energies (see Table III). Second, the differences arising in the sub-band-edge energies as obtained from different boundary conditions also depend on the amount of mixing occurring in the bulk wave functions in the [001] direction. Depending on the change in the material parameters (such as E_g, Δ, etc.), across the interface, the wave-function mixing can be shown to be no more than a few percent (see below). Third, the TMM requires an initial wave function at one side of the interface. Once this is given, the wave function is obtained fairly accurately at all points in a layer. The exponentiation procedure used in obtaining the transfer matrix is numerically stable and the errors in the procedure are less than 1 part in 10^8. A numerical error of essentially the same order occurs in the transfer-matrix calculation for each point z, $d_i < z < d_{i+1}$, in a layer. Hence, unlike "shooting" and other discretization methods used in solving differential equations, a slight error introduced in the calculation of the wave function through the use of boundary conditions across an interface occurs only at the interface region. This error is transferred to the other regions only as a shift of the value of the wave function which appears throughout the layer under consideration. The error is local and does not accumulate, and hence its influence on the subband energies continues to be of second order in perturbation theory.

TABLE III. Subband energy levels (in meV) for $Ga_{0.47}In_{0.53}As$/InP rectangular wells for the boundary conditions (BC) discussed in the text.

d (Å)	Bands	BC1	BC2	BC3
15	c	211.46	211.46	212.06
	lh	224.16	224.13	249.08
	hh	129.43	129.43	165.10
30	c	150.25	150.25	150.44
	lh	138.79	138.76	153.47
	hh	56.19	56.19	68.72
		220.14	220.14	246.63
47	c	102.35	102.36	102.16
	lh	89.70	89.69	97.89
		257.38	257.36	270.50
	hh	28.18	28.18	32.75
		112.43	112.43	127.20
		248.33	248.33	270.05
79	c	57.72	57.73	57.45
		196.92	196.92	197.23
	lh	49.06	49.05	52.63
		158.51	158.49	165.93
		263.82	263.81	271.70
	hh	11.95	11.95	13.21
		47.82	47.82	52.49
		107.45	107.45	116.6
		190.04	190.04	202.59
		292.08	292.08	302.35

Our explanations can be further illustrated by comparing Tables I and III. The momentum matrix element of $Ga_{0.47}In_{0.53}As$ is almost 23% larger than that of InP, while the momentum matrix element of GaAs is about 3% larger than that of $Ga_{0.63}Al_{0.37}As$. Also $Ga_{0.47}In_{0.53}As$ has a smaller band gap than that of GaAs. Thus band mixing in the $Ga_{0.47}In_{0.53}As$/InP QW is larger than in the GaAs/$Ga_{0.63}Al_{0.37}As$ QW. As an example, for rectangular well of 47-Å QW's, the light holes at their second sub-band-edge energy in GaAs/$Ga_{0.63}Al_{0.37}As$ have about 0.5% wave mixing with spin-orbit split-off holes, while the corresponding mixing in $Ga_{0.47}In_{0.53}As$/InP is about 2.3%. Thus the calculation

for $Ga_{0.47}In_{0.53}As/InP$ heterostructures should be more sensitive to the applied boundary conditions than that for $GaAs/Ga_{0.63}Al_{0.37}As$, according to the above arguments. This is why the difference between the results obtained from BC1 and BC2 in the case of $Ga_{0.47}In_{0.53}As/InP$ is larger than that in the case of $GaAs/Ga_{0.63}Al_{0.37}As$. For the same reason, one expects that the calculation for conduction-sub-band-edge energies should be less sensitive to the boundary conditions than those for light holes. From Table I, the results from the diagonal TMM are in excellent agreement with previous TMM calculations.[16] The slight differences are due to the cutoff technique used to prevent numerical overflow[16] in the previous development of the TMM. Such cutoffs are no longer needed here.

We list the allowed band-to-band transition energies from our calculation with the theoretical and experimental values and their uncertainties of exciton transition energies from Ref. 21 in Table IV. The corrections of exciton binding energy are also considered. These exciton binding energy corrections are taken from Lin *et al.*,[25] who have used a thermally modulated photoluminescence technique[26] to study systemically the exciton binding energy in the $Ga_{0.47}In_{0.53}As/InP$ quantum well. These exciton binding energies are also in close agreement with the analysis from the dips beside the 1*c*-1hh response peaks observed in the photoluminescence excitation spectrum for the 30- and 47-Å quantum wells.[21] Our theoretical results are in reasonable agreement with the experimental data. But the splittings between sub-band-edge energies of light holes and heavy holes are generally smaller than the experimental values. This may be attributed to fluctuations in In concentration in growing those $Ga_{0.47}In_{0.53}As/InP$ QW samples. A richer In concentration in the sample introduces biaxial compression in the well material, which can increase the valence-sub-band splitting. We fit the experimental data using In concentration and well width as two adjustable parameters. The fitting ranges were limited to the uncertainties of the experimental characterization of the samples. The results for transition energies are much improved by these adjustments. The only exception is that for the 47-Å well where a good fitting can only be obtained at In concentration at about 57.2%. The 2*c*-2hh and 2*c*-2lh transitions observed in that sample cannot be reproduced theoretically with the best-known conduction-band offset ratio,[27] unless we increase the conduction-band offset ratio to 52% and In concentration to 58%. We calculate the transition energies for 1*c*-1hh, 1*c*-1lh, 2*c*-2hh, and 2*c*-2lh transitions. The theoretical values are 0.913, 1.008, 1.198, and 1.369 eV, to be compared with the experimental values of 0.903, 1.018, 1.198, and 1.373 eV, respectively. (This adjustment, however, does not improve matters

TABLE V. Material parameters for $Ga_{1-x}Al_xAs$. The fractional conduction-band offset is assumed to be independent of composition x and a value of 0.7 is used (Ref. 28).

Parameter	$Ga_{1-x}Al_xAs$
E_g (eV)[a]	$1.5192+1.452x-0.9x^2+1.1x^3$
Δ (eV)	$0.341-0.1282x+0.0685x^2-0.016x^3$
E_p (eV)[b]	$22.7138-1.6138x$
F[c]	0.0
γ_1[b]	$1.806-0.131x$
γ_2[b]	$-0.5676+0.2241x$

[a]Reference 30.
[b]Linear interpolation from corresponding data of GaAs and AlAs in Ref. 5.
[c]The effect from remote conduction bands to the effective mass of electrons is neglected.

TABLE IV. Comparison of transition energies between experiment and theory for $Ga_{0.47}In_{0.53}As/InP$ rectangular wells. The theoretical estimates in columns 4–6 are those after excitonic corrections.

d (Å)	Transition	Expt.[a]	Theory[a]	Theory[b]	Best fit[c]	$\bar{d}$(Å)[d]	$\bar{x}$[e]
14.67	1*c*-1hh	1.130±0.010	1.151	1.137	1.131	14.67	0.542
	1*c*-1lh	1.252±0.022	1.254	1.248	1.246		
29.35	1*c*-1hh	1.013±0.010	1.006	1.010	1.016	26.41	0.550
	1*c*-1lh	1.141±0.024	1.104	1.101	1.122		
46.95	1*c*-1hh	0.907±0.010	0.929	0.935	0.912	46.59	0.572
	1*c*-1lh	1.018±0.022	0.995	1.004	1.007		
79	1*c*-1hh	0.858±0.005	0.870	0.875	0.857	85.10	0.546
	1*c*-1lh	0.905±0.010	0.908	0.913	0.905		
	2*c*-2hh	1.018±0.015	1.050	1.047	1.026		
	2*c*-2lh	1.145±0.020	1.173	1.153	1.140		

[a]Data read from Ref. 21. The reading error is $\sim\pm2$ meV. The uncertainties in the experimental values are taken from their error bars.
[b]Calculated values using either BC1 or BC2.
[c]The input parameters for $Ga_{1-x}In_xAs$ around $x=0.532$ are discussed in the text.
[d]The well width d is fitted by varying d by monolayer thicknesses within the experimental uncertainty.
[e]The In concentration $\bar{x}$ is fitted by varying it within the experimental uncertainty.

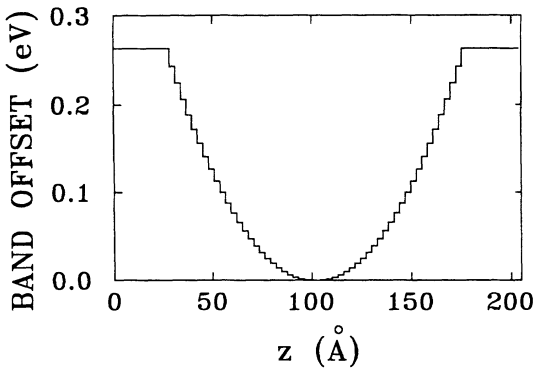

FIG. 1. The conduction-band offset as a function of coordinate in a 147-Å $Ga_{1-x}Al_xAs/Ga_{0.7}Al_{0.3}As$ parabolic quantum well.

TABLE VI. Subband energy levels (in meV) for the 147-Å $Ga_{1-x}Al_xAs/Ga_{0.7}Al_{0.3}As$ parabolic well for the boundary conditions discussed in the text. The last column presents the corresponding result in the one-band model with no nonparabolicity.

Bands	BC1	BC2	BC3	One-band model
c	49.39	49.39	49.39	48.94
	145.48	145.48	145.85	149.05
	231.80	231.80	231.80	240.1
lh	26.47	26.47	26.47	26.70
	77.60	77.60	77.60	80.50
hh	14.59	14.59	14.59	14.59
	44.17	44.17	44.17	44.17
	73.70	73.70	73.70	73.70
	101.38	101.38	101.38	101.38

for the other QW's.) The best fitting parameters as well as their results are listed in Table IV. The presence of strain due to any lattice mismatch introduced by our changes in the In concentration has been accounted for in our calculations. We should mention that we have neglected the contribution of remote conduction bands to the effective mass of the electron in the model; it has been suggested that this may not be able to adequately describe $Ga_{0.47}In_{0.53}As$ material.[22] This may account for some deviations between theory and experiment.

In the final example we present a calculation for a parabolic QW to demonstrate the power of the diagonal TMM in obtaining energy levels in heterostructures with arbitrary band-edge profiles in a multiband model. This example allows us to explore the case where the band-edge profile varies continuously in a heterostructure. Theoretically such a system has boundary conditions only at infinity since in effect there are no interfaces. However, an analytic solution cannot be obtained in this case. We approximate the parabolic well by thin layers of homogeneous composition. The band-edge profile for a 147-Å wide parabolic QW with a discretization of 50 layers is shown in Fig. 1. The well is made of $Ga_{1-x}Al_xAs$, with the Al concentration x varying from 0 to 0.3, and the barrier is made of $Ga_{0.7}Al_{0.3}As$. The material parameters for the layers are shown in Table V. The conduction-band-offset of the band-gap mismatch is assumed to be independent of composition.

The results for the energy levels of carriers in the parabolic QW are listed in Table VI. These results are entirely independent of the applied boundary conditions. For example, the results obtained from BC1 are the same as from the others. This is not the case in the rectangular quantum well. One can view the material parameters on either side of an imaginary interface to be essentially equal in this case. Thus, regardless of the form of the boundary conditions, the energy levels obtained are the same when the band edge varies smoothly. We also list the result of calculations from a one-band model without including any band nonparabolicity for reference. The first electron subband energy is lower than that obtained from the multiband model, and subsequent subband energies are higher. This is to be expected from the effect of nonparabolicity in QW.[7] The first subband energy of electrons calculated without considering the band nonparabolicity in the parabolic QW is about 3 meV lower than that obtained from the formula

$$2(n-\tfrac{1}{2})\hbar\omega\ , \quad (n=1,2,\ldots)\ ,$$

where $\hbar\omega$ is the ground-state quantum confinement energy for an infinite parabolic well.[28] This discrepancy reflects the finite depth of the parabolic well as well as a heavier effective mass of the electron in such a parabolic quantum well.

In conclusion, we have presented a unified TMM algorithm for the subband energy levels and wave functions for both SL's and QW's with any arbitrary band-edge profile in a multiband model. We have also demonstrated numerically that the effect of the boundary condition on the determination of sub-band-edge energies for wide-band-gap heterostructures in multiband models are much less important than is conventionally believed.

ACKNOWLEDGMENTS

Two of us (B. C. and M. L.) wish to thank Professor D. F. Nelson for encouragement. Their work was supported by Army Research Office contract No. DAAL03-89-K-0030.

[1]G. Bastard, Phys. Rev. B **24**, 5693 (1981); **25**, 7584 (1982); in *Molecular Beam Epitaxy and Heterostructures*, edited by L. L. Chang and K. Ploog (Nijhoff, Dordrecht, 1985), p. 381.

[2]S. R. White and L. J. Sham, Phys. Rev. Lett. **47**, 879 (1981).

[3]J. M. Luttinger and W. Kohn, Phys. Rev. **97**, 869 (1955); J. M. Luttinger, *ibid.* **102**, 1030 (1956).

[4]E. O. Kane, in *Semiconductors and Semimetals*, edited by R. K. Willardson and A. C. Beer (Academic New York, 1966),

Vol. 1, p. 75.
[5]*Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology,* edited by O. Madelung Group III (Springer, Berlin, 1982), Vol. 17.
[6]R. de L. Kronig and W. G. Penney, Proc. R. Soc. London Ser. A **130**, 499 (1931).
[7]D. F. Nelson, R. C. Miller, and D. A. Kleinman, Phys. Rev. B **35**, 7770 (1987).
[8]M. Altarelli, Phys. Rev. B **28**, 842 (1983); M. Altarelli, U. Ekengberg, and A. Fasolino, *ibid.* **32**, 5138 (1985).
[9]L. C. Andreani, A. Pasquarello, and F. Bassani, Phys. Rev. B **36**, 5887 (1987).
[10]W. Pötz, W. Porod, and D. K. Ferry, Phys. Rev. B **32**, 3868 (1985).
[11]R. Eppenga, M. F. H. Schuurmans, and S. Colak, Phys. Rev. B **36**, 1554 (1987).
[12]M. G. Burt, Semicond. Sci. Technol. **3**, 739 (1988).
[13]C. Mailhiot, T. C. Mcgill, and D. L. Smith, J. Vac. Sci. Technol. B **2**, 371 (1984).
[14]D. L. Smith and C. Mailhiot, Phys. Rev. B **33**, 8345 (1986); Rev. Mod. Phys. **62**, 173 (1990).
[15]E. P. O'Reilly and G. P. Witchlow, Phys. Rev. B **34**, 6030 (1986).
[16]L. R. Ram-Mohan, K. H. Yoo, and R. L. Aggarwal, Phys. Rev. B **38**, 6151 (1988).
[17]J. M. Berroir, Y. Guldner, J. P. Vieren, M. Voos, and J. P. Faurie, Phys. Rev. B **34**, 891 (1986).
[18]K. H. Yoo, R. L. Aggarwal, and L. R. Ram-Mohan, J. Vac. Sci. Technol. A **7**, 415 (1989); J. R. Meyer, R. J. Wagner, F. J. Bartoli, C. A. Hoffman, and L. R. Ram-Mohan, Phys. Rev. B **40**, 1388 (1989); K. H. Yoo, R. L. Aggarwal, L. R. Ram-Mohan, and O. K. Wu, J. Vac. Sci. Technol. A **8**, 1194 (1990); M. Dobrowolska, T. Wojtowicz, H. Luo, J. K. Furdyna, O. K. Wu, J. R. Meyer, C. A. Hoffman, F. J. Bartoli, and L. R. Ram-Mohan, Semicond. Sci. Technol. **5**, s103 (1990); C. A. Hoffman, J. R. Meyer, R. J. Wagner, F. J. Bartoli, X. Chu, J. P. Faurie, L. R. Ram-Mohan, and H. Xie, J. Vac. Sci. Technol. A **8**, 1200 (1990); J. R. Meyer, F. J. Bartoli, C. A. Hoffman, and L. R. Ram-Mohan, Phys. Rev. B **42**, 9050 (1990); M. Dobrowolska, T. Wojtiowicz, J. K. Furdyna, J. R. Meyer, R. D. Feldman, R. F. Austin, and L. R. Ram-Mohan, Appl. Phys. Lett. **57**, 1781 (1990).
[19]C. Moler and C. van Loan, SIAM Rev. **20**, 801 (1978).
[20]K. H. Yoo, L. R. Ram-Mohan, and D. F. Nelson, Phys. Rev. B **39**, 12 808 (1989).
[21]D. Gershoni, H. Temkin, and M. B. Panish, Phys. Rev. B **38**, 7870 (1988).
[22]K. Alavi and R. L. Aggarwal, Phys. Rev. B **21**, 1311 (1980).
[23]O. Berolo and J. C. Woolley, in *Proceedings of the Eleventh International Conference* (Polish Scientific, Warsaw, 1972), p. 1420.
[24]T. P. Pearsall, in *GaInAsP Alloy Semiconductor,* edited by T. P. Pearsall (Wiley, New York, 1982), p. 295.
[25]Z. H. Lin, T. Y. Wang, G. B. Stringfellow, and P. C. Taylor, Phys. Lett. **52**, 1590 (1988).
[26]M. Gal, C. P. Kuo, B. Lee, R. Ranganathan, P. C. Taylor and G. B. Stringfellow, Phys. Rev. B **34**, 1356 (1986).
[27]D. V. Lang, M. B. Panish, F. Capasso, J. Allam, R. A. Hamm, A. M. Sergent, and W. T. Tsang, Appl. Phys. Lett. **736**, (1987).
[28]R. C. Miller, A. C. Gossard, D. A. Kleinman, and O. Munteanu, Phys. Rev. B **29**, 3470 (1984).
[29]D. F. Nelson, R. C. Miller, C. W. Tu, and S. K. Sputz, Phys. Rev. B **36**, 8063 (1987).
[30]R. Dingle, R. A. Logan, and J. T. Arthur, Jr., in *Proceedings of the Sixth International Symposium on Gallium Arsenide and Related Compounds, Edinburgh, 1976,* IOP Conf. Ser. No. 33a, edited by C. Hilsum (IOP, Bristol, 1977), Chap. 4, pp. 210–214.